REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
30/04/2009	Final	February 1, 2006-January 31, 2009

4. TITLE AND SUBTITLE

Continued Development of the Look-up-Table (LUT) Methodology for Interpretation of Remotely Sensed Ocean Color

5a. CONTRACT NUMBER

5b. GRANT NUMBER

N00014-06-1-0370

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

W. Paul Bissett, Ph.D.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Florida Environmental Research Institute
PO Box 292397
Tampa, Florida 33687-2397

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
One Liberty Center
875 North Randolph Street, Suite 1425
Arlington, VA 22203-1995

10. SPONSOR/MONITOR'S ACRONYM(S)

ONR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for Public Release; Distribution is Unlimited

13. SUPPLEMENTARY NOTES

N/A

14. ABSTRACT

See Attached

15. SUBJECT TERMS

Remote sensing, colored dissolved organic matter, CDOM, inherent optical properties, IOP's, bathymetry, HyperSpectral Imaging, HSI, Look-Up-Table, LUT

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE		2	Beverly Walters
					19b. TELEPHONE NUMBER *(Include area code)* 800-928-6402, ext. 101

The problem of extracting environmental information from remotely sensed ocean color spectra is fundamental to a wide range of Navy needs as well as to basic science and ecosystem monitoring and management problems. Extraction of bathymetry and bottom classification is especially valuable for planning military operations in denied access areas. The ability to simultaneously generate error estimates on retrieved values is often equally important to the ability to retrieve the environmental information itself; this can be accomplished using the kNN techniques reported in this project.

We have developed mechanisms that (1) provide a more robust prediction than the previous LUT methods, and (2) provide confidence intervals to the predictions. We focus on a k-Nearest Neighbor (kNN) approach where we select not the single best match, but rather the top 50 matches. We provide a statistical measure that allows us to describe the range around the predicted estimate of bathymetry. The kNN approach does produce a more robust, accurate map of bathymetry than using a single value LUT approach. In addition, the ability to use spatial correlations to filter anomalous values and improve retrievals, as developed in this project, greatly enhances the reliability of the retrievals.

Continued Development of the Look-up-table (LUT) Methodology For Interpretation of Remotely Sensed Ocean Color Data

W. Paul Bissett
Florida Environmental Research Institute
PO Box 292397
Tampa, FL 33687-2397 USA

phone: (800) 928-6402 x102 fax: (800) 928-6402 email: pbissett [at] flenvironmental [dot] org

Award Number: N000140610370
http://www.FERIweb.org
http://www.onr.navy.mil/sci_tech/32/322/ocean_optics_biology.asp

LONG-TERM GOAL

The overall goal of this work is to refine and validate a spectrum-matching and look-up-table (LUT) technique for rapidly and accurately inverting remotely sensed hyperspectral reflectances to extract environmental information such as water-column optical properties, bathymetry, and bottom classification.

OBJECTIVES

We (PI - Bissett, and C. Mobley, N0001406C0177) are developing and evaluating techniques for the extraction of environmental information including water-column inherent optical properties (IOPs) and shallow-water bathymetry and bottom classification from remotely-sensed hyperspectral ocean-color spectra. We address the need for rapid, automated interpretation of hyperspectral imagery. The research issues center on development and evaluation of spectrum-matching algorithms, including the generation of confidence metrics for the retrieved information.

APPROACH – Year 1

In previous work, a Look-Up Table (LUT) algorithm was used in accurately predicting bathymetry (Mobley et al. 2002, Bissett et al. 2004, Bissett et al. 2005, Mobley et al. 2005). In this study we apply the k-Nearest Neighbor (kNN) algorithm. Instead of basing a prediction upon a single nearest neighbor, as in the case of the LUT, a prediction is arrived at via a combination of multiple entries within the table. We show that this method leads to an increase in the accuracy and precision of bathymetry. The additional predicted values allow for the creation of confidence intervals around the mean. This would not be possible using only the single nearest neighbor. We show that the confidence interval surrounding the observed mean is also likely to contain the true depth. The kNN algorithm can also be used on bottom type and inherent optical properties. When the classes are categorical rather than real-valued, the majority vote is used in order to determine the appropriatc classification.

When classifying new spectra, the Euclidean distance between each spectrum in the database is calculated. The k nearest neighbors to that spectra (those having the smallest distances), are

considered sufficiently qualified to predict the corresponding class labels. In the general case, the class of a pixel is determined by a majority vote from the k nearest neighbors. In the event of a tie, a prediction is made randomly from amongst the majority classes. This method is equivalent to the standard LUT algorithm when k=1.

In the event the prediction is a real-valued number, such as the case with bathymetry, additional information is available by considering the predictions as a random variable having a measurable distribution. In our case, the mean prediction from all of the k neighbors is calculated. When using the mean rather than a majority vote, predictions with smaller granularity than the LUT itself can be made. Consider the case where half of the closest entries predict a depth of 5.0m and the other half predict a depth of 5.5m. Instead of deciding randomly between an equal majority of 5.0m and 5.5m predictions, a more reasonable prediction might be 5.25m.

A nearest neighbor algorithm which produces predictions based solely upon the closest matching pattern is prone to noise. This noise is inherent in the training set. The spectrum for one particular depth, bottom type, and set of inherent optical properties may closely match a multitude of spectra from a few different classes. kNN relies on the increased probability that a spectrum presented for classification will come from the majority class of proximate training spectra, rather than a single "lucky" spectrum. A two dimensional representation of this problem is shown in Figure 1. While the target example "square" is closest in distance to the training example "O", it is more likely that the correct prediction is "X".

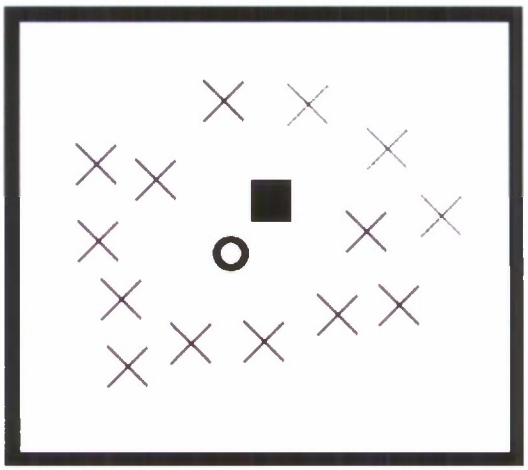


Figure 1. Xs and Os are the classes of examples belonging to the training set. The target concept, □, is closer to the O than any X. In kNN, multiple nearest neighbors vote on the appropriate class. If k = 1, class O is chosen. If k > 1, class X is chosen, the number of which is dependent on the value of k, and in which will include O in the retrieved set. The estimate of the target concept may then be calculated from any number of statistical calculations on the set of Xs, e.g. mean, majority vote, etc.

By using multiple nearest entries in the LUT, the manifestations of sensor and atmospheric noise in the predicted image can also be reduced. Small but random fluctuations of the test pixels may result in different predicted values, resulting in a "speckling" of the image. In contrast, choosing the majority class creates a less variable space from which to make a decision, making it is less likely to produce different classifications due to small amounts of noise in the spectra.

Real valued predictions enable the calculation of an observed mean, $\hat{\mu}$. They also allow for the calculation of an interval surrounding $\hat{\mu}$ by recording the observed standard deviation, $\hat{\sigma}$. We characterize the predicted depth values as being distributed as a random variable $X \sim N(\mu, \sigma^2)$ and build a confidence interval around $\hat{\mu}$ such that with probability $p = 1 - \alpha$, $\hat{\mu}$ is contained within such an interval. This method does not provide a probability that the true depth is actually contained within such an interval (it may be that the training set does not include the appropriate classes to produce an accurate match). Rather, we hope that the true depth is also contained within this interval.

Let n be the number of neighbors used in calculating the mean. From Statistics, Equation (1) establishes bounds around $\hat{\mu}$.

$$\begin{aligned} 1-\alpha &= P(-z \le Z \le z) \\ &= P(-z \le \frac{\hat{\mu}-\mu}{\hat{\sigma}/\sqrt{n}} \le z) \\ &= P(\hat{\mu} - \frac{z\hat{\sigma}}{\sqrt{n}} \le \mu \le \hat{\mu} + \frac{z\hat{\sigma}}{\sqrt{n}}) \end{aligned} \qquad (1)$$

Unfortunately, the number of training examples used is not particularly helpful in calculating the standard error $\hat{\sigma}/n$. In particular, the presumption of independence in (1) is particularly problematic since a computer generated table of spectra is simply being resampled n times without replacement in obtaining the value of $\hat{\mu}$. The number of different samples containing any one depth value is limited by the parameters used during generation of the LUT. For this reason, we have chosen to set n equal to one regardless of the number of neighbors used. This simplification is shown in Equation (2).

$$\hat{\mu} - z\hat{\sigma} \le \mu \le \hat{\mu} + z\hat{\sigma} \qquad (2)$$

Finally, the LUT may not allow for depth predictions no finer than half of the specified granularity between depths, g. While it is possible for $\hat{\mu}$ to be any real number, this is not reliable (all predicted depths could equal 1.0m when the true depth is 1.1m). This necessitates the addition of granularity into the equation. Our final confidence interval is shown in Equation 3.

$$\hat{\mu} - Z\hat{\sigma} - \frac{1}{2}g \le TrueDepth \le \hat{\mu} + Z\hat{\sigma} + \frac{1}{2}g \qquad (3)$$

We construct a database of Rrs spectra using Ecolight. In generating the database, 25 IOPs, 122 bottom types, and 50 depth values ranging from 0.01m to 25.0m in 0.5m increments were used. The corresponding Rrs spectra were generated for each combination of these three variables. The total number of entries in the LUT was 25 * 122 * 50 = 152,500; each entry containing 40 wavelengths from .4025 to .5975 microns in .005 micron intervals.

We have applied these algorithms to an image of Looe Key (Figure 2). The Looe Key image is a 3060x3257 image with 1,977,774 pixels available for classification. Ground truth data is available for the bathymetry, obtained via LIDAR. This allows additional analysis to be performed on the accuracy of our bathymetry predictions, and on the validity of our confidence interval.

Figure 2. This HSI was collected during October 2002 during a joint NRL, NAVO, FERI, USACE exercise to collect simultaneous LIDAR and hyperspectral data. The data used in this study is from the red region shown in the center of the image.

We use a z value of 3 in generating the confidence bounds. This corresponds to an α value of .001349. We choose k=50 nearest neighbors which is equal to the number of depth values for each iop and bottom type combination in the table.

RESULTS – Year 1

In Figure 3, the true bathymetry map is shown with the predicted bathymetry map using only the single closest neighbor and using the 50 nearest neighbors. Visually, using the 50 nearest neighbors produces a smoother and more lifelike view of the bathymetry than using only a single neighbor. Comparisons with the depths given in the LIDAR image show that using the additional entries in the LUT for classification and averaging the results produces a more accurate representation of the true depth.

Figure 3. Depth maps for k=1 nearest neighbors (top) vs k=50 nearest neighbors (middle) and the LIDAR data (bottom). Using 50 nearest neighbors produces a smoother more accurate graph when compared against 1 nearest neighbor to the LIDAR data.

The mean squared error using one entry from the LUT is 2.84 meters, meanwhile it is 2.60 meters using 50 entries. This represents a reduction in error by approximately 8.5%. Two histograms showing the frequency of the errors in meters are shown in Figure 4. Using 50 nearest neighbors produces two peaks each closer to 0m error than using only 1 neighbor. Likewise, errors of greater than -7.5m are no longer present in the data, and using 50 entries clearly reduces the largest errors of approximately -8.75m. As the two histograms are nearly overlapping for points 0 to 0.5m in error, using the larger number of neighbors is doing more than simply increasing depth classifications. Notice also that while the algorithm peaks at zero, very rarely are depths shallower than predicted.

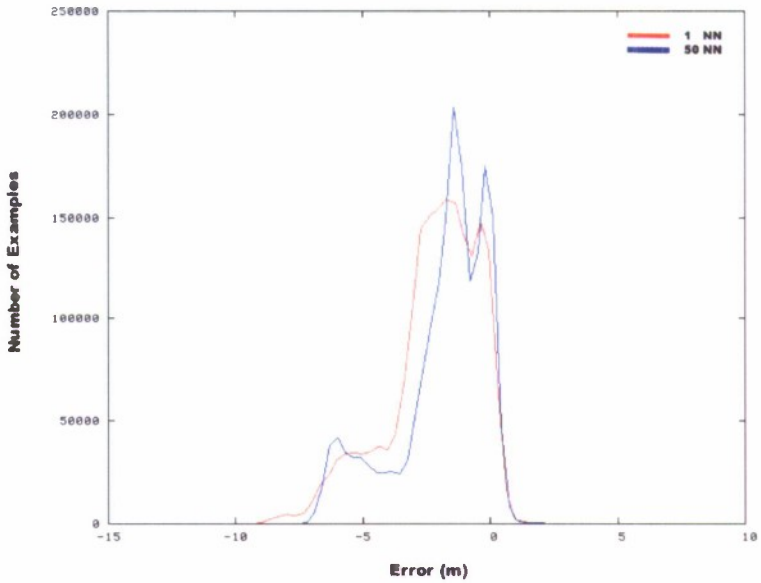


Figure 4. Depth Error Histogram for k=1 nearest neighbor and k=50 nearest neighbors. Negative values have been predicted as being shallower than the true depth.

For this problem, the availability of LIDAR ground-truth data allows us to locate which pixels fall outside of the range of the error bounds. This map is shown in Figure 5. The large area marked with a box where the confidence interval predominantly fails to accurately contain the true mean is Hawk Channel. The water in this location has been previously identified as having large errors resulting from IOP classes that were not described in the training set. This highlights the errors found when a measured spectra is produced by classes beyond those available to the training set, and in this case is beyond the recognition capabilities of the kNN algorithm.

An interesting observation is made by comparing the ratio of the margin of error to the mean with the percentage of error for each pixel. In other words, we compare the value d in $Depth \pm d\%$ to the error percentage after classification. This comparison is found in Figure 6. Namely, an image of these d values looks very similar to an inverted image of the percentage of error map. This implies that low d values correspond *negatively* with low error, and vice versa. One would expect that a low margin of error would mean the algorithm is very confident in its decision, and the percentage of error would be

very low. For this image, this turns out to be almost exactly the opposite.

Analysis has shown that shallow depths produce a much higher signal-to-noise ratio (μ/σ) than do their deeper counterparts. Yet we know from experimentation that more accurate predictions are made at shallow depths. There are two separate factors causing this seemingly non-intuitive interaction. First, the even granularity of the training set means that a prediction which is "one-off" affects shallower depths greater than it does deeper depths. For example, in our dataset, a pixel for which kNN is repeatedly predicting 1.5m and 2.0m will have a higher standard deviation to mean percentage (also known as coefficient of variance) than repeatedly predicting 24.5 and 25.0m.

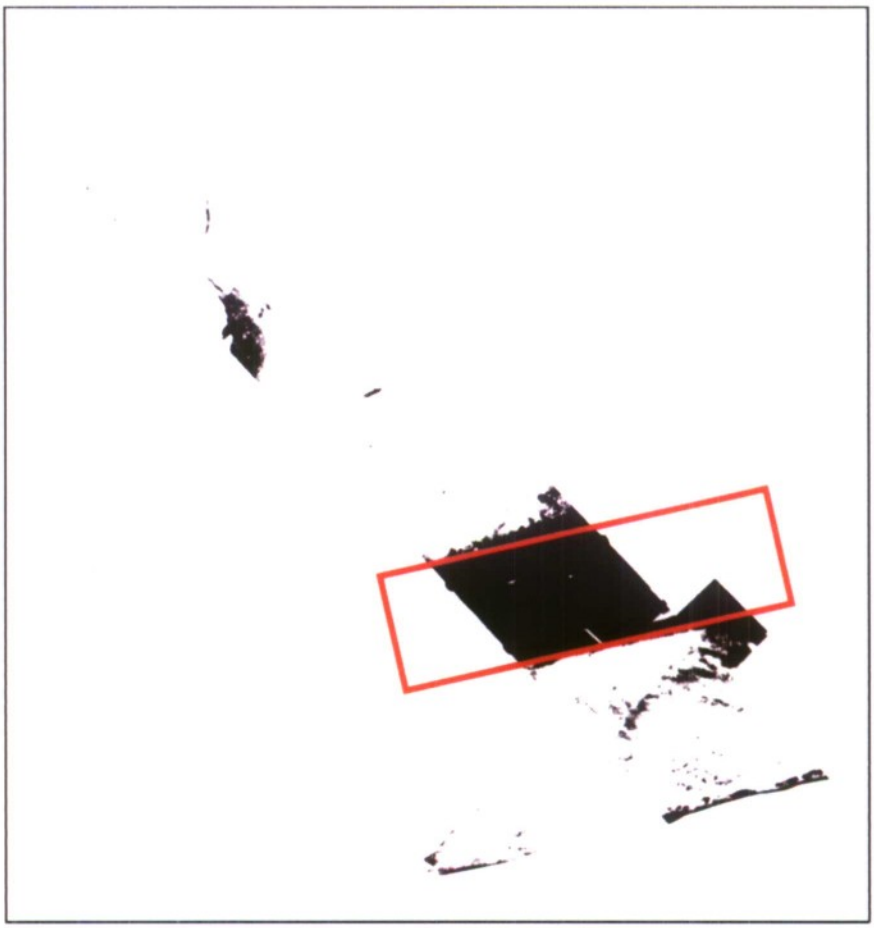


Figure 5. Error interval accuracy map. Black pixels exceeded the error bounds. The red box surrounds Hawk Channel, a region of very deep water beyond the recognition capability of this algorithm.

Furthermore, deep pixels have less coloration due to bottom type than do shallower pixels. Assuming the closest spectra in the training set is in fact the spectra containing the appropriate depth, bottom type, and IOPs, the next closest spectra is necessarily going to belong to a different one of any or all of these classes. Ideally all the correct depths would be identified as neighbors. However at shallow depths, bottom types have a greater affect on the spectra, hence neighbors are more likely to be of the same bottom type, not depth. Hence, the neighbors are more stable at deeper depths.

SUMMARY – Year 1

The previous use of the LUT approach provided good agreement between predicted and measured values of bathymetry. We have been experimenting with mechanisms that (1) provide a more robust prediction to the real-value and categorical classes, and (2) provide confidence intervals to the predictions. In this study, we focus on a k-Nearest Neighbor (kNN) approach where we select not the single best match, but rather the top 50 matches. We provide a statistical measure that allows us to describe the range around the predicted estimate of bathymetry in which we would expect that the true bathymetry to lie. The selection of a larger number of predictions does not ensure that we have an accurate prediction, for if the measured remote sensing reflectance spectrum is a function of classes (bathymetry, IOPs, and bottom reflectance) not contained with the training set, then all the predictions will be in error (see Hawk Channel example). However, the kNN approach does produce a more robust, accurate map of bathymetry that using a single value LUT approach.

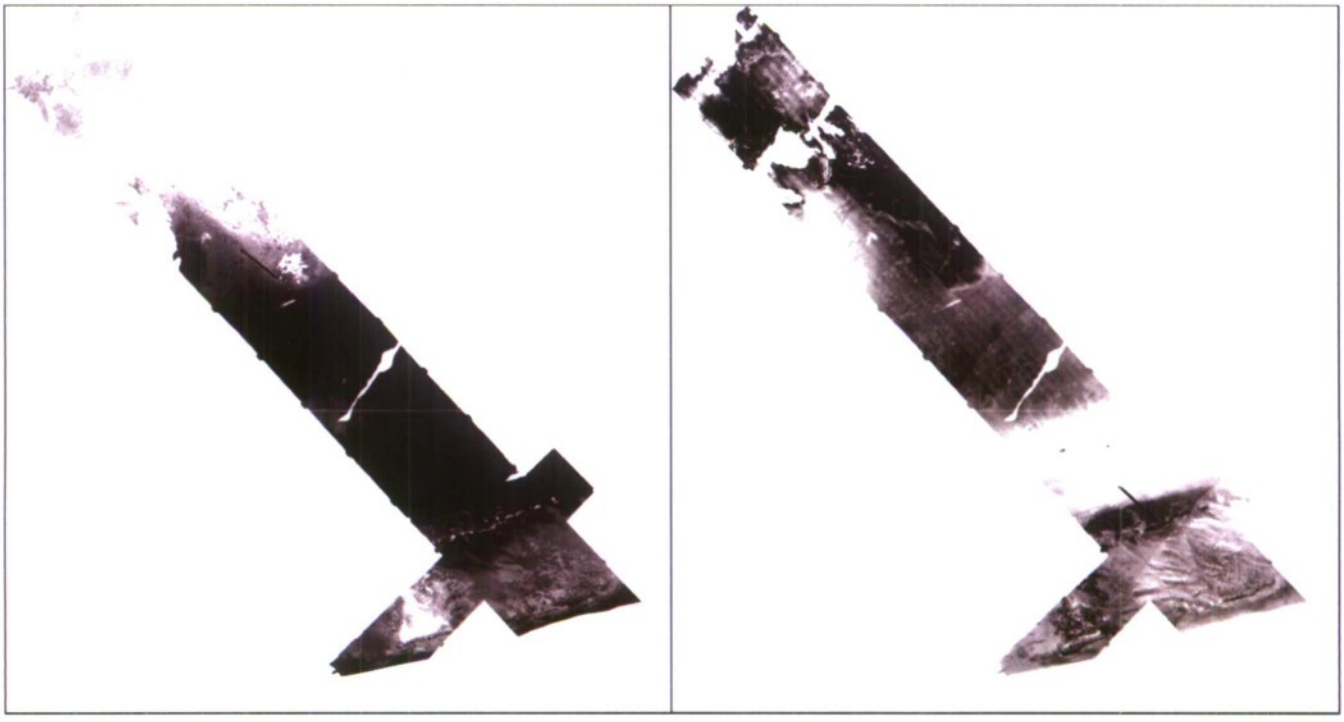


Figure 6. In the confidence map on the left, whiter pixels have a larger margin of error. In the percent difference map on the right, whiter pixels have greater error. These seems counter intuitive until (1) the relationship between granularity and depth is considered (shallower depths are more impacted by the errors induced from coarse data bathymetry), and (2) the deeper waters will exhibit more similarity given that the spectra will be less impacted by bottom reflectance, thereby yield a lower range in the estimated bathymetry.

APPROACH – Year 2

The LUT methodology is based on a spectrum-matching and look-up-table approach in which the measured remote-sensing reflectance spectrum is compared with a large database of spectra corresponding to known water, bottom, and external environmental conditions. The water and bottom conditions of the water body where the spectrum was measured are then taken to be the same as the

conditions corresponding to the database spectrum that most closely matches the measured spectrum.

In previous LUT work, we have simultaneously retrieved water column IOPs, bottom depth, and bottom classification at each pixel from the remote-sensing reflectance R_{rs} spectra. This is much to ask from a simple R_{rs} spectrum, but we have shown that all of this information is uniquely contained in hyperspectral reflectance signatures and that the information can be extracted with considerable accuracy (Mobley et al., 2005).

Previous work has considered only retrievals based on the closest matching LUT database R_{rs} spectrum to a given image spectrum. However, exactly which database spectrum most closely matches the image spectrum can be influenced by noise in the image spectrum. Another way to do the retrievals is to find not just the closest-fitting database spectrum, but to find the k closest fitting spectra. Each of these k spectra corresponds to different environmental conditions (bottom depth, bottom type, or water IOPs). The retrieval can then be taken as the mean value (or some other statistic, such as the most frequently occurring value) of the k values. If these k spectra all correspond to very nearly the same environmental conditions, then we can be confident that the retrieval is not strongly influenced by noise and is, presumably, correct to within a small error. However, if the k closest spectra correspond to widely differing environmental conditions, then we are much less confident of the correctness of the retrieval. A measure of the confidence in a depth retrieval can be based on the standard deviation of the distribution of the k retrieved depths, for example. It should be noted that even if the value of an environmental parameter as obtained from the k closest-matching spectra analysis value is the same as the value obtained for the closest-matching spectrum, the distribution of the k values can be used to compute error estimates for the retrieved quantity. Indeed, the real value of this technique often lies in the generation of confidence bounds on retrieved quantities, which in application may be as important at the retrieved value itself. In the literature this approach to classification and error estimation is known as k Nearest Neighbor (kNN) analysis.

WORK COMPLETED – Year 2

This year's work centered on evaluating the kNN method for obtaining quantitative measures of the uncertainty of the depth retrievals, i.e., for putting error bars on the retrieved depths at each pixel. Examples of depth retrievals and associated error metrics are shown below. We also evaluated a number of different metrics for determining the closeness of two spectra.

In addition to the work discussed here, we performed a detailed analysis of LUT depth and bottom classification retrievals in the localized area of Horseshoe Reef, Lee Stocking Island, Bahamas, for which bottom classification information was available from underwater transects by divers. The LUT results were in good agreement with ground truth for percent coverages of sediments, corals, and mixed bottom types over the reef. A paper on that work is now in press (Lesser and Mobley, in press).

We also applied the LUT methodology to imagery of optically deep turbid waters in Puget Sound, Washington. That work (not shown here) showed the need for improved methods of atmospheric correction of hyperspectral imagery because the retrievals for a given image were sensitive to the atmospheric correction scheme used (empirical line fit or TAFKKA). It was also found that in optically deep waters the LUT-retrieved bottom depth was roughly equal to the penetration depth at the wavelengths where the water was clearest, rather than a retrieval of infinitely deep water. (The penetration depth at a given wavelength is defined as the inverse of the diffuse attenuation coefficient for downwelling plane irradiance and gives an estimate of how far a sensor can "see" into the water

column at that wavelength.) The reason that LUT retrieved the penetration depth rather an infinite depth in optically deep waters is not yet understood.

During this period we also performed an analysis of several statistical measures of "best" fit of the kNN retrieved bathymetry estimates using the 2002 FERI/NAVO/NRL/USACE Joint Looe Key HyperSpectral Imaging (HSI) and LIDAR Experiment (Bissett et al, 2005; Figure 7). This analysis included both vector distance and angular separation in an attempt to determine which measure of best fit would be appropriate for use in retrieving bathymetry, IOPs, and bottom classification estimates.

RESULTS – Year 2

The LUT approach to retrieving IOPs, bottom reflectance, and bottom depth information from remote-sensing reflectances has performed well in its application to various PHILLS images of optically clear and shallow waters (e.g., Mobley, et al., 2005). This year we re-analyzed imagery from the Lee Stocking Island (LSI), Bahamas, area, for which acoustic bathymetry data were available to study the utility of kNN analysis in generating error maps corresponding to the retrieved bathymetry maps. Figure 7 shows an RGB PHILLS image taken near LSI; Figure 8 shows the corresponding acoustic bathymetry.

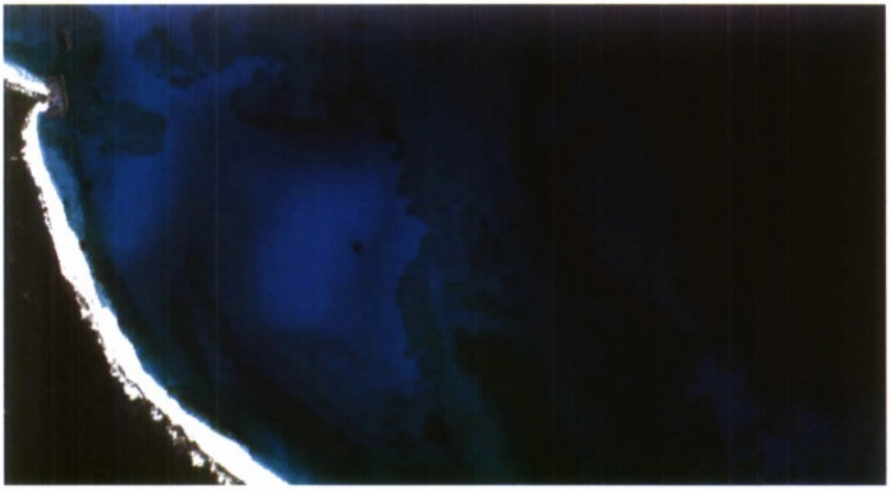


Figure 7. An RGB image of the Horseshoe Reef area made from a PHILLS hyperspectral image taken May 20, 2000. The bottom includes areas of highly reflecting ooid sands, low reflecting, dense sea grass beds, and low to intermediate reflecting areas of mixed sediments, corals, sea grass, turf algae, and macrophytes.

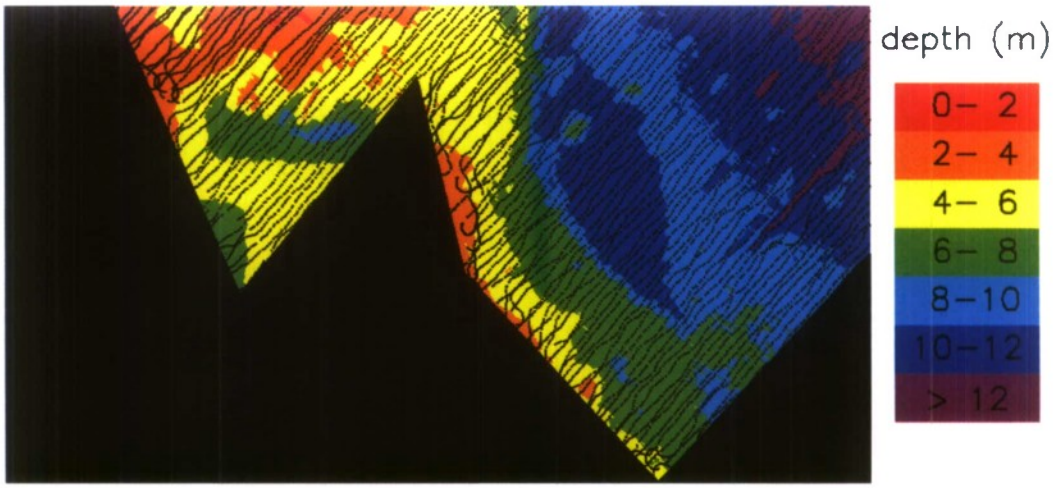


Figure 8. Acoustic bathymetry coverage for the area corresponding to Fig. 1. The black dots show the locations of the acoustic pings; the solid black area had no acoustic coverage. The acoustic depths are used for validation of the LUT-retrieved depths at the corresponding pixels. [The color coding identifies the depth, binned into 2 m bins for convenient viewing.]

When doing a depth retrieval on this image with $k = 1$, i.e., when using only the closest-matching database spectrum (with the Eucledian distance metric; see Table 1) at each image pixel, the LUT bathymetry was on average 7.0% or 0.4 m too shallow; 66% of the pixels were within ±1 m of the correct (acoustic) depth, and 87% of the pixels were within ±25% of the correct depth. When the retrievals were done with $k = 30$ and the retrieved depth was taken to be the mean of the 30 values, the LUT bathymetry was on average only 1.8% or 0.04 m too shallow. The other two statistics changed very little. Thus the kNN retrievals were on average deeper, which is correct, but the spread of retrieved vs. acoustic depths was essentially unchanged. This is likely because that spread of values is influenced by errors in geolocation of image vs. acoustic points (discussed in Mobley et al., 2005), which cannot be rectified by any analysis technique. Figure 9 shows the retrieved depths for $k = 1$, and Figure 10 shows the retrievals defined as the mean of 30 values. The most noticeable difference is that the deeper waters at the upper right of the image are somewhat deeper for the $k = 30$ retrieval.

The k retrieved depths at each image pixel were used to generate two kinds of error maps for the depth retrievals. Figure 11 shows the map of the standard deviation of the 30 retrieved depths. This map gives an estimate of the absolute error in the retrieved depths. As would be expected, the standard deviation of the retrieved depths is greatest for the deepest water. However, some shallow areas with dark bottoms also have large standard deviations.

Figure 12 shows the map of the standard deviation divided by the mean depth, which gives a map of the relative errors in the depth retrievals. Overall for this image, this error metric is in the 0.05 to 0.15 range, although some areas with dark bottoms have larger relative errors. In general, areas with bright bottoms (ooid sands, in this image) have the smallest relative errors in the depth retrievals.

These error maps are in qualitative agreement with what is expected from signal-to-noise considerations, i.e., bathymetry for shallow or bright-bottom areas is retrieved most accurately, and deeper areas and areas with darker bottoms have more uncertainty in the retrieved bathymetry. However, the combination of LUT spectrum matching and kNN analysis allows us to generate

quantitative error estimates on the retrieved bathymetry. Such error maps cannot be generated by simpler algorithms that work only with the image spectrum (e.g., band ratio algorithms).

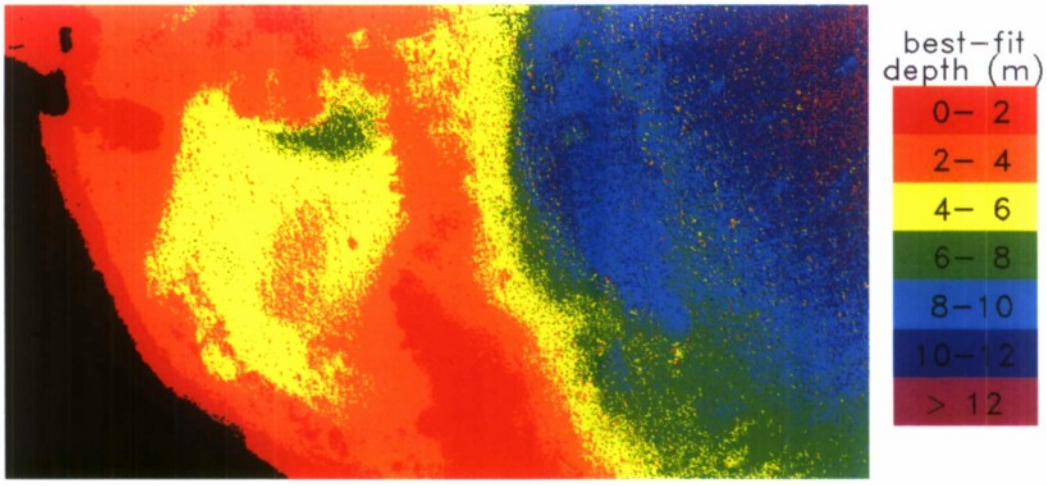


Figure 9. LUT depth retrieval obtained from the closest matching database spectrum (k = 1). [The color coding identifies the depth, binned into 2 m bins for convenient viewing.]

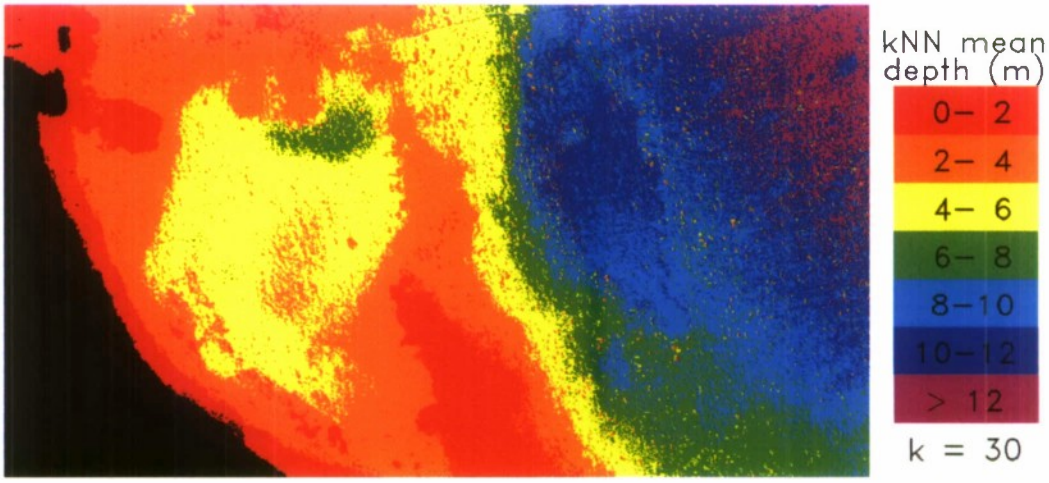


Figure 10. LUT depth retrieval computed as the mean of the k = 30 closest matching database spectra. Note that the retrievals are somewhat deeper than those for k = 1. [The color coding identifies the depth, binned into 2 m bins for convenient viewing.]

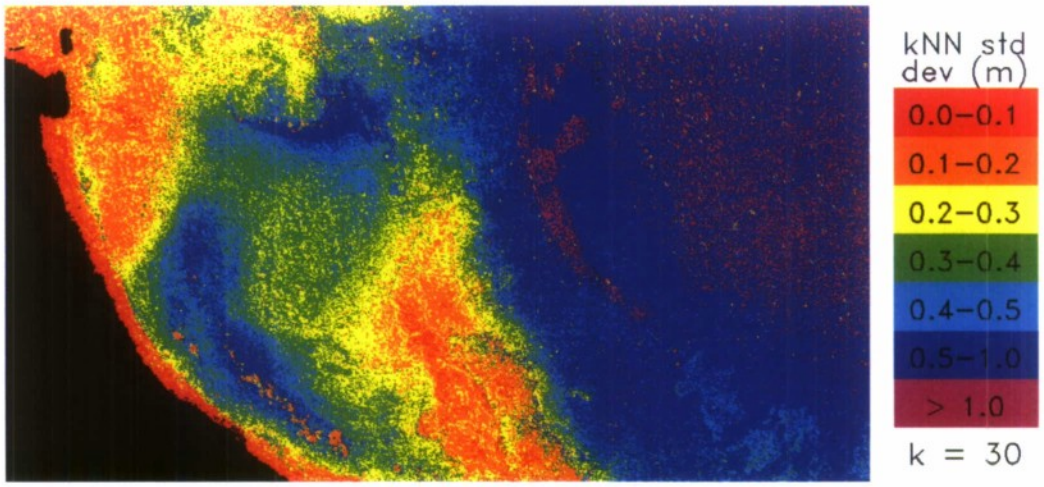


Figure 11. Standard deviation of the LUT depths for the k = 30 closest matching spectra. The standard deviations of the retrieved depths are greatest in deeper waters and in areas with dark bottoms.

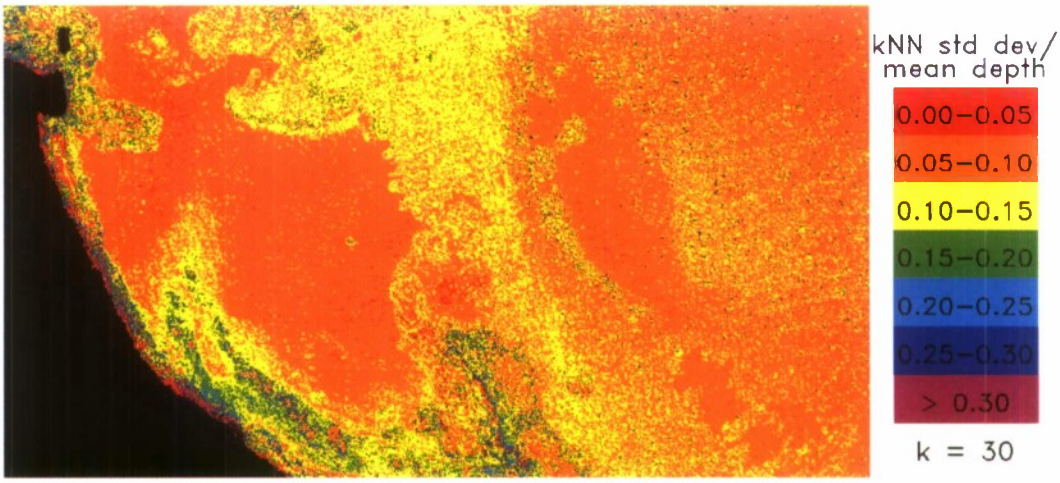


Figure 12. Ratio of the standard deviation (Fig. 11) to the mean depth (Fig. 10) for k = 30. This figure shows that the relative depth error is generally in the 0.05 to 0.15 range over most of the image. Greater relative errors occur over some areas of darker bottoms.

The analysis on Horseshoe Reef was completed using a traditional Euclidean distance calculation (see Table 1). However, there are a number of different statistical measures of "best" fit that could be used to complete the kNN retrievals. Table 1 lists those that we explored using the Looe Key HSI/LIDAR data (Figure 13).

Table 1. Metrics for measuring the closeness of two spectra i and j. Here x_{ik} means R_{rs} spectrum i at wavelength k, and the sums are over wavelength. The two spectra that give the minimum value of the metric are the closest.

Vector Distance Separation	Equation
Euclidean: Sum (over wavelength) of squared point distances	$\sqrt{\sum_{k=1}^{n}(x_{ik}-x_{jk})^2}$
Manhattan: Sum of absolute point distances	$\sum_{k=1}^{n}\lvert x_{ik}-x_{jk}\rvert$
Chebyshev: Closest absolute maximum point distance	$\max_k \lvert x_{ik}-x_{jk}\rvert$
Canberra: Sum of absolute point distances divided by absolute point values	$\sum_{k=1}^{n}\frac{\lvert x_{ik}-x_{jk}\rvert}{\lvert x_{ik}\rvert+\lvert x_{jk}\rvert}$
Bray Curtis: Sum of absolute point distances divided by sum of absolute point values	$\frac{\sum_{k=1}^{n}\lvert x_{ik}-x_{jk}\rvert}{\sum_{k=1}^{n}(x_{ik}+x_{jk})}$

Vector Angle Separation	Equation
Angular Separation: Cosine angle between two vectors	$\frac{\sum_{k=1}^{n}x_{ik}\cdot x_{jk}}{\sqrt{\sum_{k=1}^{n}x_{ik}^2\cdot\sum_{k=1}^{n}x_{jk}^2}}$
Correlation Coefficient: Cosine angle between two vectors where the coordinates are centered at the mean	$\frac{\sum_{k=1}^{n}(x_{ik}-\bar{x}_i)\cdot(x_{jk}-\bar{x}_j)}{\sqrt{\sum_{k=1}^{n}(x_{ik}-\bar{x}_i)^2\cdot\sum_{k=1}^{n}(x_{jk}-\bar{x}_j)^2}}$

Figure 14 shows the results over the joint HSI/LIDAR coverage area. It is evident in this data set that vector distance is the best measure for goodness of fit, and that the Manhattan distance calculation is the best calculation. There were some areas that were fit best by angular separation. These areas were primarily best fit by the Correlation Coefficient. The distance measure takes into consideration the magnitude of the R_{rs} signal, as well as the spectral shape. Its use is only appropriate where there is high confidence in the calibration of the sensor measured water-leaving radiance and the subsequent removal of atmospheric and illumination effects. Most spectrum matching techniques avoid the use of the magnitude because of problems in sensor calibration and atmospheric correction. We demonstrate here the enhanced retrieval resulting from the distance measurements possible with R_{rs} data using spectra retrieved from a high confidence sensor and processing.

Figure 15 shows the best fit using only Manhattan and Correlation Coefficient measures. It can be seen that there are areas where the Correlation Coefficient achieves better matches than the Manhattan. This is primarily seen in shallow water areas where the magnitude of the signal is strong enough to allow for the secondary effects of spectral shape to provide a more enhanced fit. Table 2 summarizes these results.

Table 2. Depth errors for various metrics, averaged over the entire image. RMSE = Root Mean Squared Error (meters), ABSE = Absolute Mean Squared Error (meters).

- Manhattan
  - RMSE = 2.25
  - ABSE = 1.79
- Combination Man+Cor
  - RMSE = 2.31
  - ABSE = 1.68
- Correlation Coefficient
  - RMSE = 5.97
  - ABSE = 4.54
- Best Possible Method
  - RMSE=2.07
  - ABSE = 1.43

Figure 13. Joint HSI/LIDAR Experiment coverage area. The red area denotes region of joint data coverage.

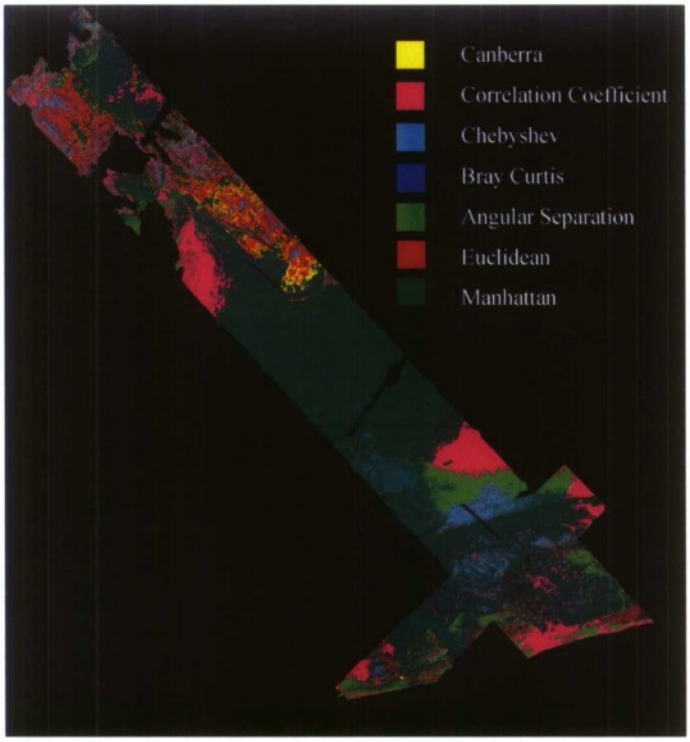


Figure 14. Visual results of statistical measure analysis of best fit. The color coding shows which distance metric gave the best agreement between LUT and LIDAR bathymetry. Overall, vector distance was better than vector angular separation in determining best fit. The Manhattan distance calculation was better than Euclidean. There were some areas where angular separation was better.

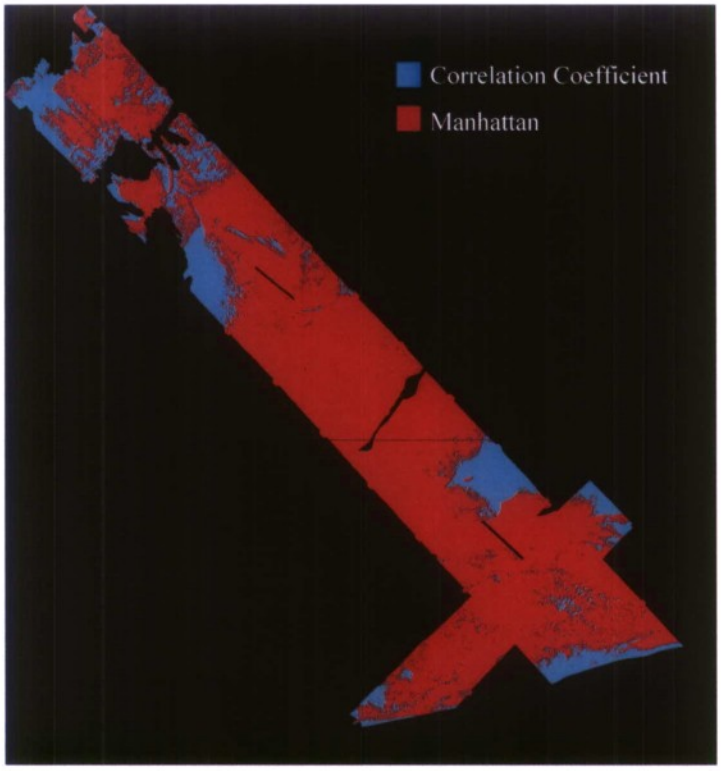


Figure 15. The same area as Figure 14, with only the best vector distance (Manhattan) and angle separation (Correlation Coefficient). This image shows that over shallow waters the Correlation Coefficient sometimes retrieved better matches than the Manhattan calculation. This results from the fact that the spectra magnitudes are large enough from the bright shallow water signal to allow for better certainty from the angular separation.

APPROACH – Year 3

The LUT methodology is based on a spectrum-matching and look-up-table approach in which the measured remote-sensing reflectance spectrum is compared with a large database of spectra corresponding to known water, bottom, and external environmental conditions. The water and bottom conditions of the water body where the spectrum was measured are then taken to be the same as the conditions corresponding to the database spectrum that most closely matches (by some chosen metric) the measured spectrum.

In previous LUT work, we have simultaneously retrieved water column IOPs, bottom depth, and bottom classification at each pixel from the remote-sensing reflectance R_{rs} spectra. This is much to ask from a simple R_{rs} spectrum, but we have shown that all of this information is uniquely contained in hyperspectral reflectance signatures and that the information can be extracted with considerable accuracy (Mobley et al., 2005; Mobley and Lesser, 2007).

Our initial work considered only retrievals based on the closest matching LUT database R_{rs} spectrum to a given image spectrum. However, exactly which database spectrum most closely matches the image spectrum can be influenced by noise in the image spectrum. Therefore, last year's work considered retrievals based not just on the closest-fitting database spectrum, but on the k closest fitting spectra. Use of the k (typically $k = 30$ to 50) closest spectra not only allows various statistical estimates of the retrieved information (depth, bottom type, etc), but also provides statistically based error bars and confidence statements about the retrieved information.

This year's work continued the process of examining various ways to improve retrievals, in particular by taking advantage of spatial correlations in the environmental variables from one pixel to the next. We also examined the errors in the LUT R_{rs} database generation associated with the use of unpolarized (scalar) radiative transfer calculations (using a special version of HydroLight), compared to exact (but very time consuming) calculations that included polarization.

WORK COMPLETED – Year 3

Previous retrievals (e.g., Mobley, et al., 2005; Lesser and Mobley, 2007) have processed each image pixel independently of its neighbors. However, there is usually a strong spatial correlation in neighboring pixels because water depth and IOPs, and bottom type, often do not change greatly from one pixel to the next (i.e., on a scale of one to a few meters). We therefore examined various ways to spatially smooth the input R_{rs} spectra (before processing) and/or the output environmental values (after processing) to take advantage of spatial similarities in small blocks of pixels (e.g., 3x3 or 5x5 blocks centered on the pixel of interest).

Colleagues Y. You, G. Kattawar, and B. Hauss and Mobley also did detailed comparison runs using coupled ocean-atmosphere vector radiative transfer codes available to You and Kattawar to quantify

the errors resulting in upwelling atmospheric radiances and in R_{rs} when the ocean and the atmosphere are modeled (1) using polarized (vector) radiative transfer (RT) theory, (2) unpolarized (scalar) RT, and (3) a vector atmosphere but a scalar ocean. We considered upwelling radiances just above the sea surface (relevant to R_{rs} spectra measured for ground truth and to the generation of the LUT database), at 3,000 m altitude (relevant to airborne remote sensing platforms as employed in our work), and at the top of the atmosphere (relevant to satellite remote sensing). A paper on those results has been submitted to *Applied Optics* (You, et al., 2008).

In addition to the work discussed above, M. Lesser and Mobely published (Lesser and Mobley, 2007) a detailed analysis of LUT depth and bottom classification retrievals in the localized area of Horseshoe Reef, Lee Stocking Island, Bahamas, for which bottom classification information was available from underwater transects by divers. The LUT results were in good agreement with ground truth for percent coverage of sediments, corals, and mixed bottom types over the reef.

Finally, a patent titled "Spectral Imaging System" was granted to the collaborators on this LUT work. That patent covers various aspects of both image acquisition hardware and image analysis software, including the LUT methodology.

RESULTS – Year 3

We investigated two types of spatial smoothing. The first smoothes the image R_{rs} spectra before performing the LUT matching, and the second smoothes the retrieved depths (or other quantities, such as IOPs or bottom reflectances) after performing the LUT matching. The two types of smoothing can be done independently or in combination, and in combination with the kNN analysis techniques investigated last year.

To spatially smooth an R_{rs} spectrum, we considered an $n \times n$ block of pixels centered on the pixel of interest, with $n = 1, 3, 5, \ldots$ ($n = 1$ corresponds to no spatial smoothing). Let $R_{rs}(i,j,\lambda)$ be the image spectrum at pixel (i,j). We reasoned that we want to average the "good" spectra in the $n \times n$ block of pixels centered on (i,j), but we do not want to include any anomalously large or small "bad" spectra that might be contaminated by sun glint or whitecaps (or other causes). For $n = 3$, we have a 3×3 block of 9 pixels centered on (i,j). To help eliminate anomalously large or small "bad" spectra, we discarded the highest and smallest values of the 9 spectra at each wavelength, and averaged the remaining 7 values. For $n = 5$, we have a 5×5 block of 25 pixels. In that case, we discard the highest 2 and lowest 2 values, and averaged the remaining 21 values. If some of the pixels are flagged as land, clouds, or whitecaps, or if (i,j) is next to the image boundary, there are fewer than n^2 valid pixels. We then have a reduced number of pixels to work with, but the procedure is the same: discard the highest and lowest values and average the remaining values. The original $R_{rs}(i,j,\lambda)$ is then replaced by the average spectrum computed from the $n \times n$ block of pixels. Note that this algorithm is applied independently at each wavelength. Thus the particular spectra that are eliminated at one wavelength may or may not be the spectra that are eliminated at another wavelength.

To smooth the retrieved depths we again consider $n \times n$ blocks of pixels. Now, however, we do not discard the high or low values of the retrieved depths before averaging. The reason is that when doing kNN matching, the kNN algorithm already, by its very nature, may have omitted the high or low values, or done some other sort of filtering or averaging of the k retrieved depths at each pixel. We therefore omit only pixels in the $n \times n$ block that are flagged as invalid (land, cloud, etc), and we then average the remaining (usually n^2) depths to obtain the spatially smoothed depth for the pixel at the

center of the $n \times n$ block.

Figure 16 shows a 3D perspective plot of the bathymetry near Lee Stocking Island (LSI), Bahamas; LSI is the gray area at the upper left of the image. This figure shows the results for no spatial smoothing ($n = 1$) of either the input R_{rs} spectra or of the retrieved depths, and the closest-matching ($k = 1$) database spectrum was used. This baseline retrieval corresponds to the retrievals shown in Mobley et al. (2005). The pixel-to-pixel variability of the retrieved depths is quite apparent. Figure 17 shows the quantitative comparison of the retrieved vs. acoustically measured depths for the pixels where an acoustic depth was available. We see that the average retrieved depth is about 7% (0.4 m) too shallow, with a standard deviation of 1.2 m between retrieved and measured depths. 68% of the pixels have retrieved depths within ±1 m of the acoustic depth, and 87% are within ±25% of the acoustic value.

Figures 18 and 19 show the corresponding results from a combination of kNN matching and spatial smoothing. Figure 15 shows the 3D plot of retrieved depths when (1) the input R_{rs} spectra are smoothed using a 5x5 spatial grid, (2) the retrieved depths are then obtained as the median of the closest $k = 30$ spectra, and (3) the retrieved depths are then smoothed using a 5x5 grid. The final retrieved depths are now clearly much smoother from pixel to pixel. Figure 19 shows the quantitative errors for Figure 18. We now see that the average error is only 0.8% (0.04 m) too shallow, with a standard deviation of 0.9 m. Now 76% of the all pixels are within ±1 m of the acoustic value, and 95% of all pixels are within ±25% of the acoustic value.

The combination of spatial smoothing and kNN analysis clearly improves the average accuracy of the retrieved depths and reduces the pixel-to-pixel variability.

The LUT R_{rs} database is generated by a special version of the HydroLight radiative transfer model, which solves the unpolarized (scalar) radiative transfer equation (RTE). Omission of polarization in the database generation leads to some error in the computed R_{rs} spectra. We therefore examined (with the assistance of Y. You, G. Kattawar, and B. Hauss) the nature of these errors through numerical simulations using coupled ocean-atmosphere radiative transfer codes that solve the polarized (vector) RTE. We evaluated the errors in upwelling radiances due to the omission of polarization in either the ocean or atmosphere for a wide range of oceanic and atmospheric conditions, sun and viewing geometries, and wavelengths from 415 to 865 nm. We considered the errors at the sea surface (relevant to R_{rs} computation for the LUT database and to sea truth measurements used for validation of remote sensing imagery), at 3,000 m altitude (relevant to airborne remote sensing platforms), and at the top of the atmosphere (relevant to satellite ocean color remote sensing).

Figures 20 and 21 illustrate the errors in the water-leaving radiance when both the ocean and the atmosphere are modeled with the unpolarized RTE (left panels), and when the atmosphere and sea-surface reflectance are polarized but the ocean is unpolarized (right panels). It is seen that the errors in the water-leaving radiance are less than 3% when the underwater RT calculations are performed using the scalar RTE (as is done in HydroLight), whether or not the atmospheric calculations are performed with scalar or vector codes. Surface waves have almost no effect on these errors compared to the effects of sun and viewing geometry and water and atmospheric conditions. These results justify the use of the computationally efficient HydroLight scalar ocean radiative transfer model to computed R_{rs} spectra for the LUT database, so long as errors up to 3% can be tolerated. This is indeed the case for most ocean color remote sensing, since the errors owing to imperfect sensor radiometric calibration

and atmospheric correction of the imagery are often greater. The results for 3,000 m and top-of-atmosphere comparisons, and for other parameter values, can be found in the You et al. (2008) paper.

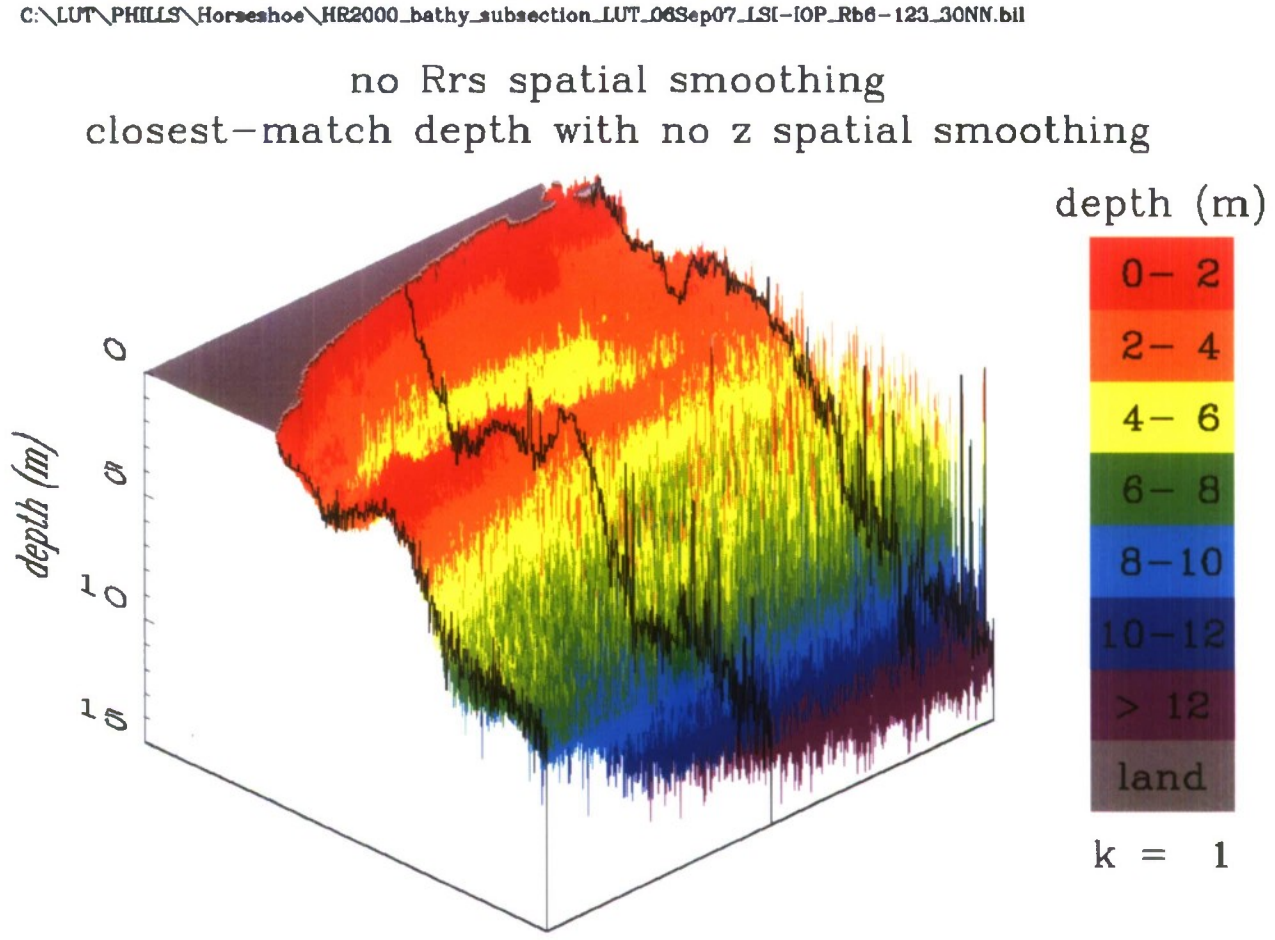


Fig. 16. LUT-retrieved depths plotted as a 3D surface and viewed in perspective. Lee Stocking Island is the gray area at the upper left. The pixel-to-pixel spikiness or variability of the depth retrievals is quite apparent. The three black lines show the depths at each pixel along selected transects of the area. The corresponding quantitative comparison of LUT-retrieved vs. acoustic-measured depths is shown in Fig. 17. [The figure shows a 3D perspective plot of retrieved depths at each pixel with the depth color coded: red is 0-2 m to purple is >12m deep.]

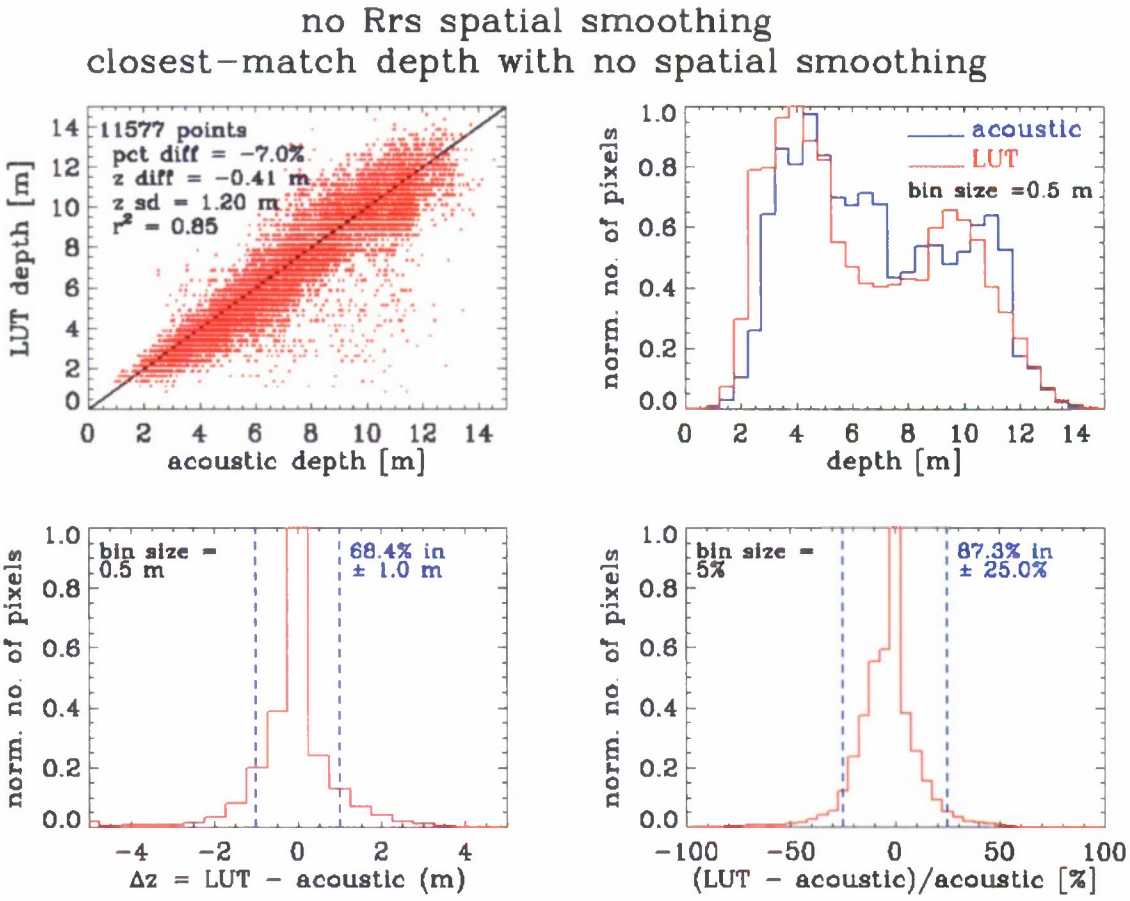


Figure 17. LUT vs. acoustic depths for the no-smoothing, closest-match retrieval of Fig. 16, displayed in various ways. [The figure compares the differences in retrieved and measured depths in four different ways.]

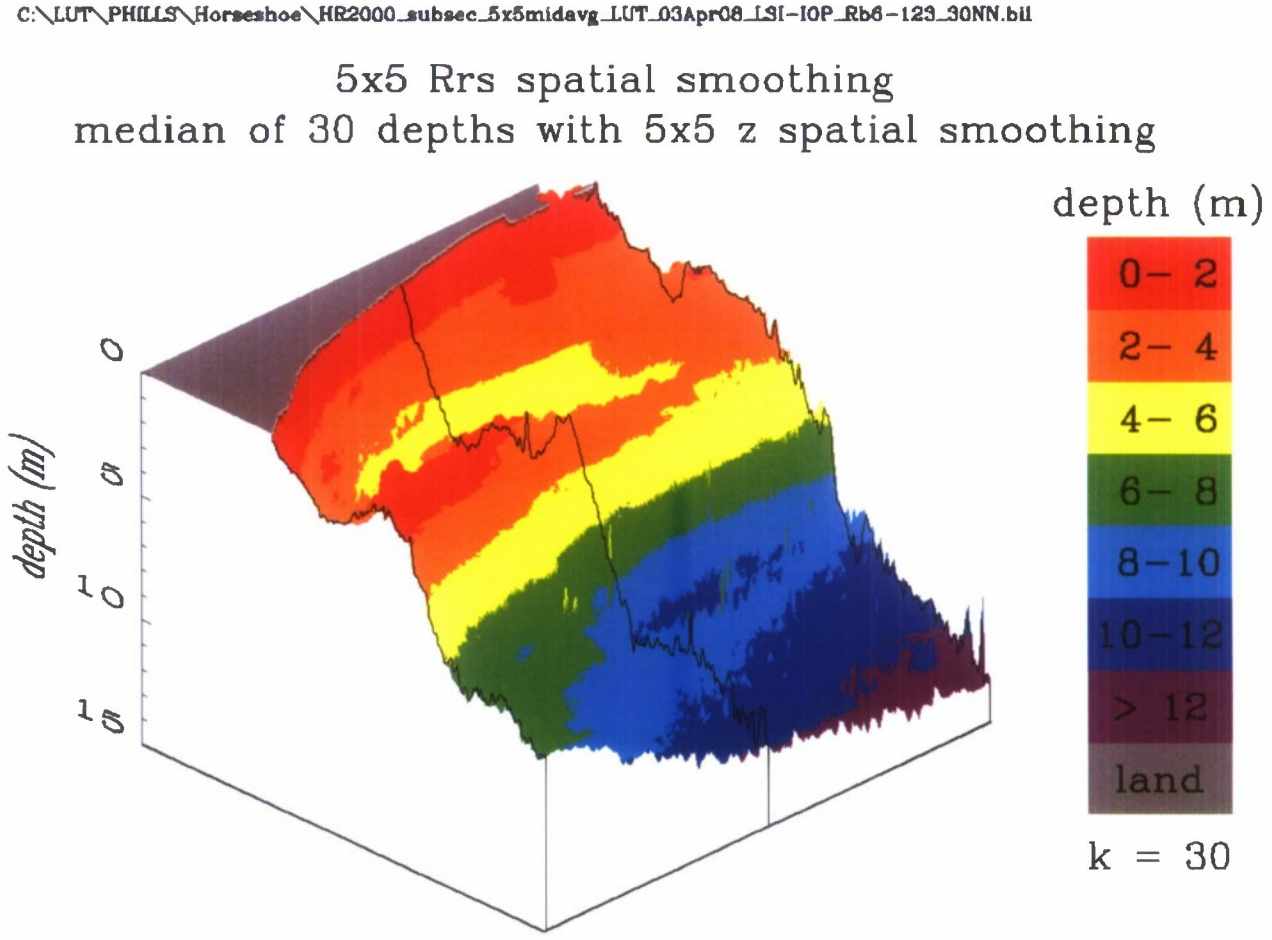
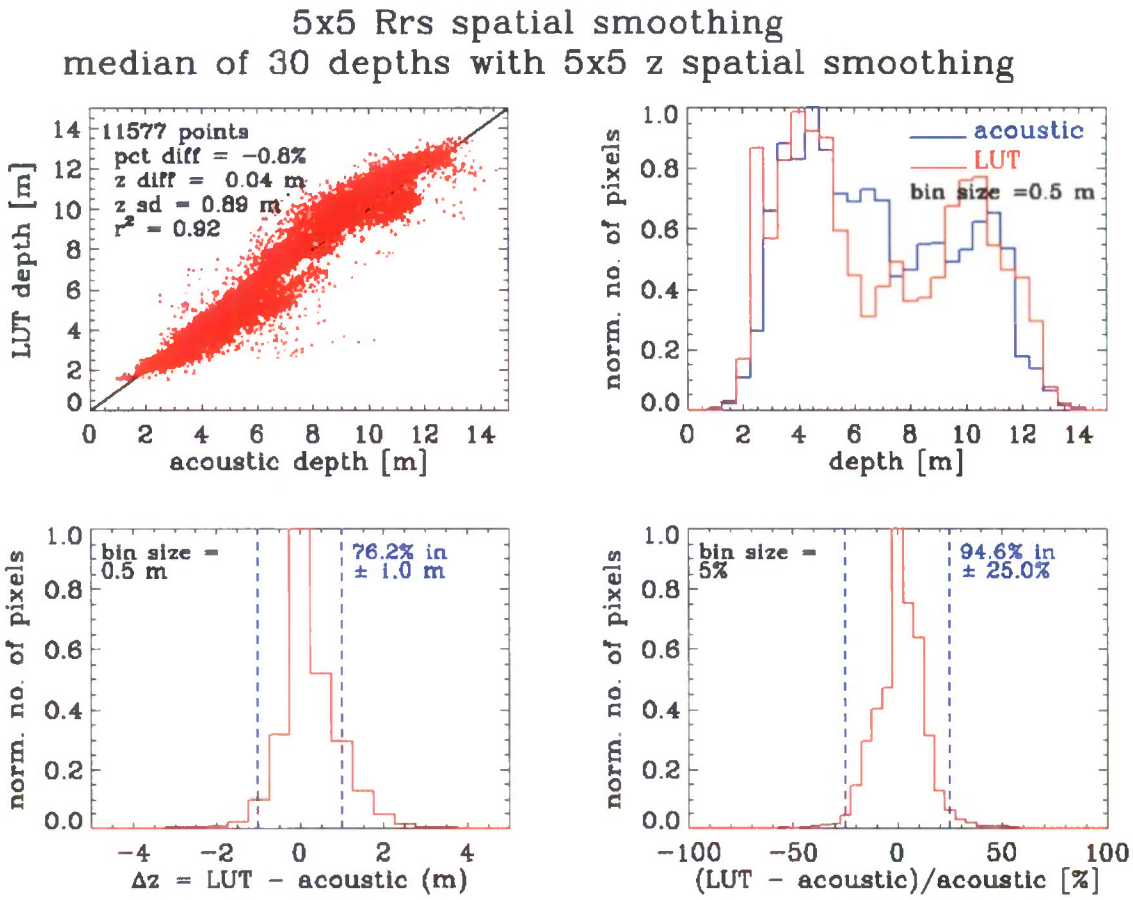


Figure 18. The 3D perspective of the retrieved depths when 5x5 spatial smoothing in performed on both the input R_{rs} spectra and the output depths, and the retrieved depth is the median of k = 30 closest matching spectra. The pixel-to-pixel spikiness or variability of the depth retrievals greatly reduced compared to what is seen in Fig. 16. The corresponding quantitative comparison of LUT-retrieved vs. acoustic-measured depths is shown in Fig. 19. [The figure shows a 3D perspective plot of retrieved depths at each pixel with the depth color coded.]



C:\LUT\PHILLS\Horseshoe\HR2000_subsec_5x5midavg_LUT_03Apr08_LSI-IOP_Rb6-123_30NN.bil
c:\lut\phills\horseshoe\acoustic_bathymetry\comp_UTM_LL_HR2000_pix.txt

Figure 19. LUT vs. acoustic depths for the smoothed kNN retrieval of Fig. 18, displayed in various ways. Compare with Fig. 17. [The figure compares the differences in retrieved and measured depths in four different ways.]

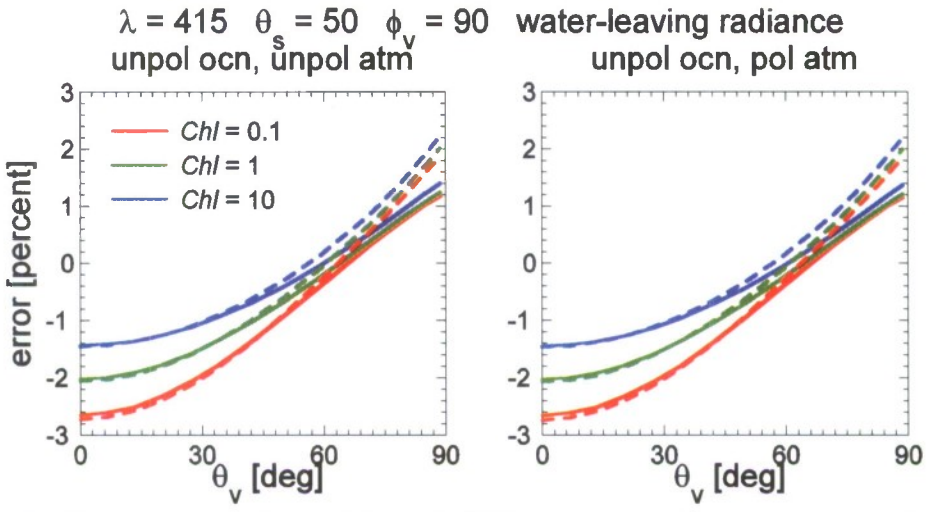


Figure 20. Errors in the water-leaving radiance just above the sea surface for low (Chl = 0.1 mg m^{-3}), medium (Chl = 1) and high (Chl = 10) chlorophyll concentrations. This figure is for 415 nm, sun at a 50 deg zenith angle, viewing at right angles to the sun, from nadir to horizon. Solid lines are for a clear atmosphere, dashed lines are for a very hazy atmosphere. [Curves show errors of less than ±3% in the water-leaving radiance for a wide range of chlorophyll concentrations.]

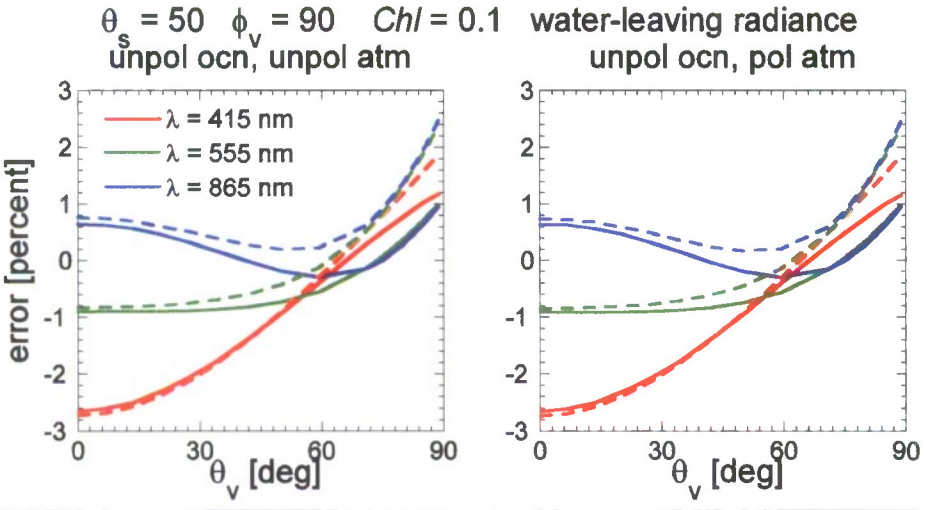


Figure 21. Errors in the water-leaving radiance just above the sea surface for wavelengths of 415, 555, and 865 nm. low (Chl = 0.1 mg m^{-3}), medium (Chl = 1) and high (Chl = 10) chlorophyll concentrations. This figure is for a low chlorophyll concentration of 0.1 mg m^{-3}, sun at a 50 deg zenith angle, viewing at right angles to the sun, from nadir to horizon. Solid lines are for a clear atmosphere, dashed lines are for a very hazy atmosphere. [Curves show errors of less than ±3% in the water-leaving radiance for a wide range of wavelengths.]

IMPACT/APPLICATION

The problem of extracting environmental information from remotely sensed ocean color spectra is fundamental to a wide range of Navy needs as well as to basic science and ecosystem monitoring and management problems. Extraction of bathymetry and bottom classification is especially valuable for planning military operations in denied access areas. The ability to simultaneously generate error estimates on retrieved values is often equally important to the ability to retrieve the environmental information itself; this can be accomplished using the kNN techniques reported in this project.

The previous use of the LUT approach provided good agreement between predicted and measured values of bathymetry. We have been experimenting with mechanisms that (1) provide a more robust prediction to the real-value and categorical classes, and (2) provide confidence intervals to the predictions. In this study, we focus on a k-Nearest Neighbor (kNN) approach where we select not the single best match, but rather the top 50 matches. We provide a statistical measure that allows us to describe the range around the predicted estimate of bathymetry in which we would expect the true bathymetry to lie. The selection of a larger number of predictions does not ensure that we have an accurate prediction, for if the measured remote sensing reflectance spectrum is a function of classes (bathymetry, IOPs, and bottom reflectance) not contained with the training set, then all the predictions will be in error (see Hawk Channel example). However, the kNN approach does produce a more robust, accurate map of bathymetry that using a single value LUT approach. In addition, the ability to use spatial correlations to filter anomalous values and improve retrievals, as developed in this project, greatly enhances the reliability of the retrievals.

TRANSITIONS

Various databases of water IOPs, bottom reflectances, and the corresponding R_{rs} spectra, along with the specialized Hydrolight code and spectrum-matching algorithms have been transitioned to other NOAA and Navy HIS projects for use in comparisons of LUT and LIDAR bathymetry.

RELATED PROJECTS

This work is being conducted in conjunction with Dr. Curtis D. Mobley at Sequoia Scientific, Inc., who is separately funded for this collaboration. These techniques developed here are now being applied to imagery of Australian coastal waters in a comparison of several different hyperspectral remote sensing algorithms for a variety of environments. That comparison study is being led by A. Dekker of CSIRO.

REFERENCES

Bissett, W.P., DeBra, S., Kadiwala, M., Kohler, D., Mobley, C., Steward, R., Weidemann, A., Davis, C.O., Lillycrop, J. and Pope, R., 2004. Development, validation, and fusion of high resolution active and passive optical imagery.Ocean Optics XVII, Fremantle, AU.

Bissett, W.P., DeBra, S., Kadiwala, M., Kohler, D.D.R., Mobley, C.D., Steward, R.G., Weidemann, A.D., Davis, C.O., Lillycrop, J. and Pope, R.L., 2005. Development, validation, and fusion of high-resolution active and passive optical imagery. In: I. Kadar (Editor), Signal Processing, Sensor Fusion, and Target Recognition XIV. Proceedings of SPIE Vol. 5809. SPIE, Bellingham, WA, pp. 341-349.

Mobley, C. D., L. K. Sundman, C. O. Davis, T. V. Downes, R. A. Leathers, M. J. Montes, J. H. Bowles, W. P. Bissett, D. D. R. Kohler, R. P. Reid, E. M. Louchard, and A. Gleason, 2005. Interpretation of hyperspectral remote-sensing imagery via spectrum matching and look-up tables. *Applied Optics* 44(17), 3576-3592.

Mobley, C.D., Sundman, L., Davis, C.O., Montes, M. and Bissett, W.P., 2002. A look-up-table approach to inverting remotely sensed ocean color data, Ocean Optics XVI. Office of Naval Research Ocean, Atmosphere, and Space S&T Department, Santa Fe, NM.

Lesser, M. P. and C. D. Mobley, 2007. Bathymetry, optical properties, and benthic classification of coral reefs using hyperspectral remote sensing imagery. *Coral Reefs*, 26: 819-829.

You, Y., G. W. Kattawar, C. D. Mobley, and B. Hauss, 2008. Polarization effects on upwelling marine atmospheric radiances. *Applied Optics* (submitted).

PUBLICATIONS

Lesser, M. P. and C. D. Mobley, 2007. Bathymetry, optical properties, and benthic classification of coral reefs using hyperspectral remote sensing imagery. *Coral Reefs*, 26: 819-829 [Refereed, published]

You, Y., G. W. Kattawar, C. D. Mobley, and B. Hauss, 2008. Polarization effects on upwelling marine atmospheric radiances. *Applied Optics* [Refereed, submitted]

PATENT

U. S. Patent 7369229 titled "Spectral Imaging System" was granted on 06 May 2008 to W. P. Bissett III, D. D. R. Kohler, R. G. Steward, and C. D. Mobley. This patent covers various aspects of both the hardward and software used for hyperspectral image acquisition and analysis. [Granted]